

Complex Adaptive Systems, Publication 3

Cihan H. Dagli, Editor in Chief

Conference Organized by Missouri University of Science and Technology
2013- Baltimore, MD

An Alternative Approach to Reduce Massive False Positives in Mammograms Using Block Variance of Local Coefficients Features and Support Vector Machine

M. P. Nguyen^{a,*}, Q. D. Truong^a, D. T. Nguyen^a, T. D. Nguyen^a, V. D. Nguyen^{a,b}^a*School of Electronics and Telecommunications, Hanoi University of Science and Technology, Hanoi, Vietnam*^b*Biomedical Electronics Center, Hanoi University of Science and Technology, Hanoi, Vietnam*

Abstract

Computer Aided Detection (CAD) systems for detecting lesions in mammograms have been investigated because the computer can improve radiologists' detection accuracy. However, the main problem encountered in the development of CAD systems is a high number of false positives usually arise. It is particularly true in mass detection. Different methods have been proposed so far for this task but the problem has not been fully solved yet. In this paper, we propose an alternative approach to perform false positive reduction in massive lesion detection. Our idea is lying in the use of Block Variation of Local Correlation Coefficients (BVLC) texture features to characterize detected masses. Then, Support Vector Machine (SVM) classifier is used to classify the detected masses. Evaluation on about 2700 RoIs (Regions of Interest) detected from Mini-MIAS database gives an accuracy of $A_z = 0.93$ (area under Receiving Operating Characteristics curve). The results show that BVLC features are effective and efficient descriptors for massive lesions in mammograms.

© 2013 The Authors. Published by Elsevier B.V. Open access under [CC BY-NC-ND license](#).

Selection and peer-review under responsibility of Missouri University of Science and Technology

Keywords: mammography; computer aided detection; false positive reduction; block variance of local coefficients; support vector machine

1. Introduction

Breast cancer is one of the most injurious and deadly diseases for women in their 40s in the United States [1] as well in the European Union [2]. More than one million breast cancer cases occur annually and more than 400,000 women die each year from this fatal disease as estimated by World Health Organization's International Agency for Research on Cancer (IARC) [3].

* Corresponding author. Tel.: + 84 946532538; fax: +84 43 8682099.

E-mail address: phuongnguyen@bme.edu.vn.

An essential point for a high survival rate in breast cancer treatment is the detection of the cancer at early stage. It is not an easy task. Mammography is a commonly used imaging modality for breast cancer to enhance the radiologists' ability to detect and diagnose cancer at an early stage and take immediate precautions for its earliest prevention [4].

The introduction of digital mammography gave the opportunity of increasing the number of commercial Computer Aided Detection (CAD) systems for detecting and diagnosing the breast cancer at an early stage [5-6]. The main reason for the mistrust of radiologists on the role of CAD system in breast cancer detection is due to a large number of false positive (FP) marks usually arises when high sensitivity is desired [7]. A FP mark is a region being normal tissue but interpreted by the CAD as a suspected one. So a CAD system for mass detection generally has a step of FP reduction.

Different approaches to reduce FPs have been proposed. Most of them are based on the extraction of features of detected suspicious regions (Regions of Interest – RoIs) such as textural features [8], geometry features [9] or Local Binary Pattern (LBP) features [10]... These features are submitted to a pattern classifier to classify the RoIs into real mass or normal parenchyma.

In this paper, we propose an alternative way to perform mass false positive reduction using moment features of the detected RoIs. Our idea is inspired by the recent work in which Block Variance of Local Correlation Coefficients (BVLC) features are applied successfully to the face recognition problem [11-12]. Once the BVLC features are extracted, Support Vector Machine (SVM) is used as pattern classifier.

We experiment the proposed method on a dataset of about 2700 RoIs that are detected from the Mini-MIAS database [13]. The obtained results demonstrate the effectiveness and efficiency of our approach. To our knowledge, this is the first attempt to use BVLC features in the field of mammographic mass detection.

2. Materials and Methods

2.1. Region of Interest detection

In this stage, suspicious regions or RoIs are extracted from the original mammogram by the CAD system. The radiologists have to focus their attention to these extracted regions. The steps of this procedure are fully described in our previous paper [14]. Detected RoIs are marked as true positive RoIs (TP-RoIs) or false positive RoIs (FP-RoIs) based on the ground truth provided in the Mini-MIAS database. There are about 2700 detected RoIs.

2.2. Feature extraction

Each detected RoI is characterized by a set of features that is formed using BVLC features. The computation of BVLC starts from correlation coefficients in a local region, which are defined as

$$\rho(x, y, r) = \frac{1}{\sigma(x, y)\sigma((x, y) + r)} \times \left[\frac{1}{|R(x, y)|} \times \sum_{(p, q) \in R(x, y)} I(p, q)I((p, q) + r) - \mu(x, y)\mu((x, y) + r) \right]$$

where r denotes a shifting orientation and $\mu(x, y)$ and $\sigma(x, y)$ are the mean and standard deviation in a local region $R(x, y)$, respectively. The terms $\mu((x, y) + r)$ and $\sigma((x, y) + r)$ are the mean and standard deviation in a local region shifted by r from (x, y) , respectively.

BVLC is then defined as

$$BVLC(x, y) = \max_{r \in O_k} [\rho(x, y, r)] - \min_{r \in O_k} [\rho(x, y, r)]$$

where O_k denotes a set of orientations with r of distance k . For instance, O_k may be chosen as $O_k = \{(-k, 0), (0, -k), (0, k), (k, 0)\}$. The value of BVLC is determined as the difference between the maximum and minimum values of the local correlation coefficients according to orientations. The higher the degree of roughness in the local region is, the larger the value of BVLC [11].

As in [12], detected regions are squared with constant size and are divided into sub regions having size of 2-by-2, 3-by-3 and 4-by-4 pixel. BVLC for each sub region is calculated then expectation and variation of BVLCs are used

as BVLC features for each region. However our detected RoIs are not squared and different from each other so we modify the calculation procedure. The new calculation procedure is as follow:

- Consider the minimal rectangular that contains the RoI.
- Divide each side of this rectangular by 2, 3 and 4. So we get 4, 9 and 16 blocks.
- Calculate BVLCs which are called BVLC2x2, BVLC3x3 and BVLC4x4 for each block.
- Similar to [12], use expectation and variation of BVLCs as BVLC features for each RoI. They are called BVLC2x2mean, BVLC2x2var, BVLC3x3mean, BVLC3x3var, BVLC4x4mean and BVLC4x4var, respectively.

So the feature set for each RoI composes of total 6 features.

An illustration of BVLC features is given in Figure 1.

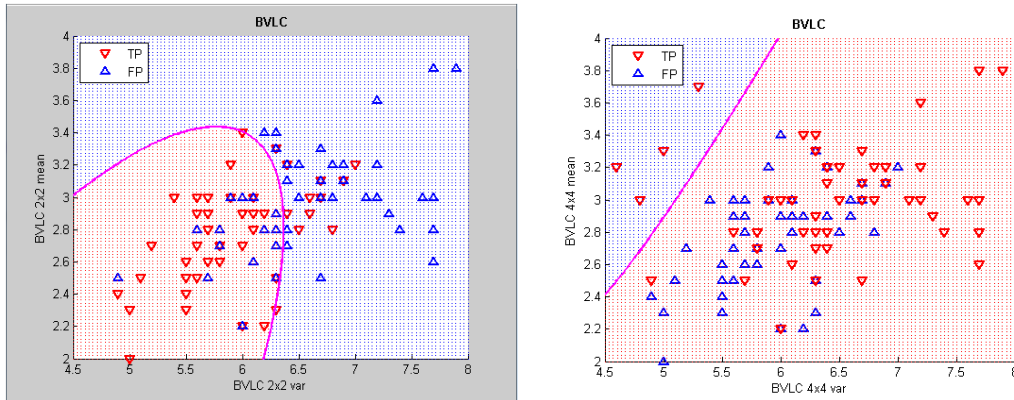


Fig. 1. (Left) BVLC2x2 feature values. (Right) BVLC4x4 feature values.

2.3. Classification

The number of features in each region are quite large bring about the number of dimension of vector space which we should consider to classify the data are sketching out. In addition, there is always the overlapping of the data class corresponding to the features in the sample region. Applying Support Vector Machine (SVM), a state-of-art classification method introduced in 1992 [15] to solve this problem takes many advantages in this case.

Given a set of training data $x_1, \dots, x_n \in \mathbb{R}^{d+1}$ with corresponding labels $y_i \in \{-1, 1\}$. SVMs seek a separating hyperplane with the maximum margin

$$\min \frac{1}{2} W^T W + C \sum_{i=1}^N \varepsilon_i \quad \text{subject to } y_i (W^T \phi(x_i) + b) \geq 1 - \varepsilon_i \text{ and } \varepsilon_i \geq 0$$

where C is the parameter controlling the trade-off between a large margin and less constrained violation. ε_i are slack variables which allow for penalized constraint violation.

Equivalently, with Lagrange multipliers $\alpha_i \geq 0$ for the first set of constraints can be used to write the optimization problem for SVMs in the dual space. By solving a quadratic programming problem, the solution for the Lagrange multiplier can be obtained. Finally, the SVM classifier takes the form

$$f(x) = \text{sign} \left(\sum_{i=1}^{\#SV} \alpha_i y_i K(x, x_i) + b \right)$$

where $\#SV$ represents the number of support vectors and the kernel function $K(.,.)$.

In this paper, we use a nonlinear SVM classifier with Gaussian RBF kernel

$$f(x) = \sum_{i=1}^m \alpha_i \cdot \exp(-\gamma \|x_i - x\|^2) + C$$

where x is input data and. α is constant while C and γ must be tuned to get the best classification performance.

2.4. Performance evaluation

To evaluate the performance of the SVM classifier, the Receiver Operating Characteristic (ROC) curve is used. The ROC curve is constructed based on two statistic factors which are the sensitivity and the specificity, and the accuracy of SVM is then computed [15]. The best possible classifier would yield when the ROC curve tends to the upper left corner representing 100% sensitivity and 100% specificity.

The accuracy value (ACC) to estimate the performance of classification process is given by

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \times 100\%$$

Another parameter is used to estimate SVM performance is area under the curve (AUC). The SVM classifier is called ideal with 100% accuracy when the AUC of its ROC approaches 1 and when AUC equals 0.5, SVM is random classifier. AUC is given by

$$AUC = \frac{\sum_{i=1}^{n^+} \sum_{j=1}^{n^-} 1_{\xi_{ij} > 0}}{n^+ n^-} \quad \text{with} \quad \xi_{ij} = f(x_i^+) - f(x_j^-)$$

where $f(\cdot)$ is denoted as decision function of classifier, x^+ and x^- respectively denote the positive and negative samples and n^+ , n^- are respectively the number of positive and negative examples and the $1_{\xi > 0}$ is defined as 1 if the predicate ξ is holds and 0 otherwise.

3. Experiments and Discussions

Our proposed method is evaluated on total number of 2700 detected RoIs [14]. Six input features as mentioned above and nonlinear SVM classifier with Gaussian RBF kernel are used. In this study, we use 10-fold cross validation method to train and test the classifier. The dataset is equally partitioned into 10 folds. For each of 10 experiments, use $(10-i)$ folds for training and i folds for testing. Each fold is used 10 times in training as well as in testing. In this evaluation, values of i is changed from 1 to 9.

Figure 2 shows obtained AUCs with different i or different ratios between training folds and testing folds corresponding to two feature subsets {BVLC2x2mean, BVLC3x3mean, BVLC4x4mean} or {BVLC Mean} and {BVLC2x2var, BVLC3x3var, BVLC4x4var} or {BVLC Var}. It is easy to realize that in most of case, using {BVLC Var} feature subset gives higher AUC value. The best AUC value is archived with $i = 6$.

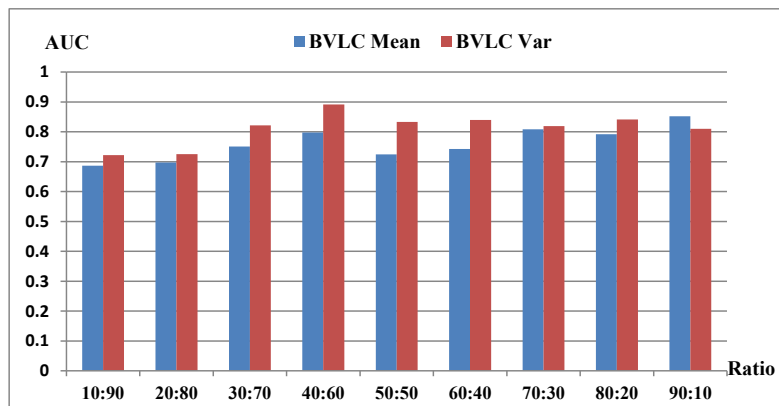


Fig. 2. AUC values with different training and testing fold ratios.

We also assess effects of different feature combinations on the performance of the SVM classifier. The results corresponding to the case $i=6$ are given in Table 1. For each type of BVLC features, BVLC mean or BVLC var,

BVLC2x2 feature always gives result better than that of BVLC3x3 or BVLC4x4. It is very clear if looking back to Figure 1. Comparisons between each pair of BVLC mean and BVLC var features indicate BVLC var features have better discrimination efficiency.

Table 1. AUC and ACC values with different feature combinations for the case $i=6$

Performances	BVLC mean			
	2x2	3x3	4x4	All
AUC	0.73903	0.7288	0.6527	0.7975
ACC (%)	86.55	83.31	72.36	88.72
	BVLC var			
	2x2	3x3	4x4	All
AUC	0.8404	0.8033	0.7955	0.8915
ACC (%)	76.34	82.64	78.25	84.67
	BVLC var + BVLC mean			
	2x2	3x3	4x4	All
AUC	0.8745	0.8015	0.6772	
	All BVLC var + BVLC2x2mean			
	2x2	3x3	4x4	All
AUC	0.9325 (± 0.0005)			

The AUC values when combining all BVLC mean or BVLC var features are 0.8915 and 0.7975, respectively. However, combining both BVLC var features and BVLC mean features does not lead to an increase in classification outcome. That fact causes us to think of combining all BVLC var features and one of BVLC mean features. Experimental results have approved our idea. We acquire the best AUC value of 0.9325 ± 0.0005 when using all BVLC var features and BVLC2x2mean feature (table 1). The value of C and γ is 1.85 and 0.0025, respectively. The corresponding ROC curves are given in Figure 3 (left). In this case, false positives reduce 82%.

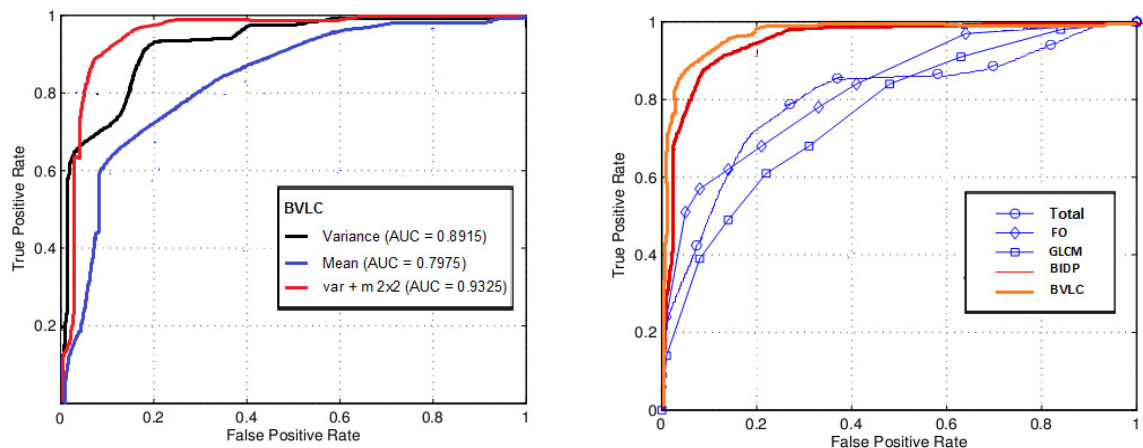


Fig. 3. (Left) ROC curves with different BVLC feature subsets. (Right) Comparison between BVLC features and FOS, BDIP, GLCM features.

The high obtained $A_z = 0.9325 \pm 0.0005$ and high false positive reduction of 82% are quite prospective. To illustrate the effectiveness of BVLC features, we compare them to other features such as FOS (First Order Statistic), GLCM (Gray Level Co-occurrence Matrix) and BDIP [11] (Block Difference Inverse Probability) features. The

AUC values are given in Table 2 and the ROC curves are illustrated in Fig 3 (right). It is obviously that using BVLC features is an effective and efficient approach to reduce false positives.

Aiming to have general trends of performance comparison, we also compare our method with other techniques on the basis of AUC value as given in Table 3. It indicates that the method we propose has a potential to be further investigated.

Table 2. Comparing BVLC features to other features for the case $i=6$

Feature type	AUC
FOS	0.6935
GLCM	0.7839
BDIP	0.9102
All BVLC var + BVLC2x2mean	0.9325

Table 3. Comparison with other methods based on AUC values

Research work	Approach	Database	No. of RoIs	AUC
A. Oliver et. al. [10]	LBP	DDSM	1024	0.906±0.043
X. Lladó et. al. [16]	LBP	DDSM	1792	0.91±0.04
B. Ioan et. al. [17]	Gabor	MIAS	322	0.79
Q. D. Truong et. al [18]	BDIP	MIAS	2700	0.9102
Proposed method	BVLC	MIAS	2700	0.9325±0.005

4. Conclusions

In this paper, we have introduced an alternative approach to reduce false positives in mammography based on BVLC features and SVM. Experiments have shown that BVLC features are effective and efficient descriptors for massive lesions in mammograms. In comparison with other descriptors, BVLC also provides better and more constant results.

In the future, combining BVLC features and other efficient features will be investigated. Selecting optimal features will be studied also.

Acknowledgment

The authors would like to thank Vietnam National Foundation for Science and Technology Development (NAFOSTED) for their financial support to publish this work.

References

1. Bray F., McCarron P., Parkin D. M., "The changing global patterns of female breast cancer incidence and mortality," *Breast Cancer Research* 6, pp. 229-239, 2004
2. Eurostat, Health statistics atlas on mortality in the European Union, Official Journal of the European Union, 2002
3. <http://globocan.iarc.fr/factsheet.asp>
4. Buseman S., Mouchawar J., Calonge N., Byers T., "Mammography screening matters for young women with breast carcinoma," *Cancer* 97(2), pp. 352–8, 2003
5. Birdwell R. L., Ikeda D. M., O'Shaughnessy K. D., Sickles E. A., "Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection," *Radiology* 219, pp. 192–202, (2001)
6. R. F. Brem, J. A. Rapelyea, G. Zisman, J. W. Hoffmeister, and M. P. DeSimio, "Evaluation of breast cancer with a computer-aided detection system by mammographic appearance and histopathology," *Cancer* 104(5), pp. 931–935, 2005

7. Taylor P., Champness J., Given-Wilson R., Johnston K., Potts H., "Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography," *Health Techn Assess* 9(6):pp. 1–58, 2005
8. M. Masotti, N. Lanconelli, and R. Campanini, "Computer aided mass detection in mammography: False positive reduction via gray scale invariant ranklet texture features," *Med. Phys.* 36(2), pp. 311-316, 2009
9. D. Tralic, J. Bozek, and S. Grgic, "Shape analysis and classification of masses in mammographic images using neural networks," in *Systems, Signals and Image Processing (IWSSIP)*, 2011 18th International Conference on. IEEE, 2011, pp. 1–5.
10. A. Oliver, X. Llado, J. Freixenet, and J. Matri, "False positive reduction in mammographic mass detection using local binary patterns," *Proceedings of the 10th international conference on Medical image computing and computer-assisted intervention - Volume Part I* Pages 286-293, 2007
11. H. J. So, M. H. Kim, and N. C. Kim, "Texture classification using wavelet-domain BDIP and BVLC features", 17th European Signal Processing Conference (EUSIPCO 2009), pp. 1117-1120, 2009
12. T. D. Nguyen, T. Q. Tran, D. T. Man, Q. T. Nguyen, M. T. Hoang, "SVM classifier based face detection system using BDIP and BVLC moments," *Proceedings of ATC2012*, 2012
13. <http://peipa.essex.ac.uk/info/mias.html>
14. V. D. Nguyen, D. T. Nguyen, H. L. Nguyen, D. H. Bui, T. D. Nguyen, "Automatic identification of massive lesions in digitalized mammograms," *Proceeding of 4th International Conference on Communications and Electronics (ICCE2012)*, 2012
15. B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers" *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, 144–152, 1992
16. Lladó X., Oliver A., Freixenet J., Martí R., Martí J., "A textural approach for mass false positive reduction in mammography," *Computerized Medical Imaging and Graphics* 33(6):415-422, 2009
17. B. Ioan, A. Gacsadi, "Directional features for automatic tumor classification of mammogram images," *Biomedical Signal Processing and Control*, Vol 6(4), pp.370-78, 2011.
18. Q. D. Truong, M. P. Nguyen, V. T. Hoang, H. T. Nguyen, D. T. Nguyen, T. D. Nguyen, V. D. Nguyen, "Feature Extraction and Support Vector Machine Based Classification for False Positive Reduction in Mammographic Images," *Proceedings of 5th International Symposium on IT in Medicine and Education (ITME2013)*, July 19-21, 2013